

**SUICÍDIO NO RN E SUA RELAÇÃO COM DETERMINANTES ESPACIAIS,
URBANIZAÇÃO, DESENVOLVIMENTO E OUTROS FATORES SOCIOECONÔMICOS***

Stefany Silva Amaral

Doutorando do Programa de Pós-Graduação em Economia da Universidade Federal de Pernambuco (PIMES/UFPE)

E-mail: s.amaral@outlook.com

RESUMO: Este trabalho trata o suicídio como um fenômeno genérico, com efeitos e causas eminentemente econômicas e sociais. Pretende-se inicialmente fazer uma breve exposição teórica e empírica de material levantado na literatura acerca do suicídio e seus determinantes espaciais e socioeconômicos e testar parte dessas relações propostas para os municípios do estado do Rio Grande do Norte, utilizando a Análise Exploratória de Dados Espaciais (AEDE) e a econometria espacial. Nos resultados foram observados *clusters* da taxa de suicídio e foi verificado o efeito contágio. Em um modelo SAR, foram encontradas relações diretas do suicídio com pessoas com ensino superior e com a desigualdade de renda e relação inversa com o nível de riqueza e a razão de dependência.

Palavras-chave: Taxa de suicídio; AEDE; Econometria espacial.

Classificação JEL: C21; I10; Z13.

**SUICIDE IN RN AND ITS RELATIONSHIP WITH SPATIAL DETERMINANTS,
URBANIZATION, DEVELOPMENT AND OTHER SOCIOECONOMIC FACTORS**

ABSTRACT: This work deals with suicide as a generic phenomenon with eminently economic and social effects and causes. The objective of this study is to present a brief theoretical and empirical exposition of the literature about suicide and its spatial and socioeconomic determinants and to test part of these relationships proposed for the municipalities of the state of Rio Grande do Norte, using Exploratory Spatial Data Analysis (ESDA) and spatial econometrics. In the results, clusters of the suicide rate were verified and the contagious effect was verified and in a SAR model, direct suicide relations were found with people with higher education and income inequality and inversely with the level of wealth and the dependency ratio.

Keywords: Suicide rate; ESDA; Spatial econometrics.

JEL Codes: C21; I10; Z13.

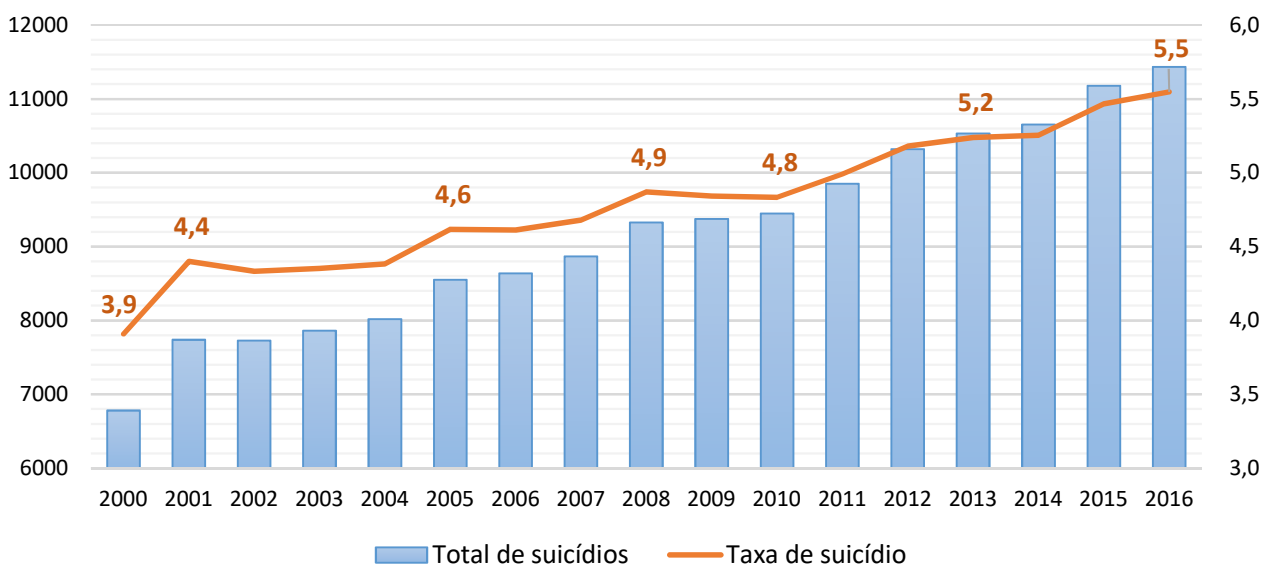
1. Introdução

O suicídio é um assunto que no meio acadêmico parece apresentar já uma larga literatura, observada não apenas na área da saúde, todavia constantemente explorado pelas ciências sociais. Em economia, a disposição de um largo ferramental empírico e profundo conhecimento das relações sociais e econômicas possibilita uma exploração produtiva dessa temática, visando à formulação de políticas e otimização da alocação de recursos da área da saúde, buscando eficiência mediante redução dos gastos e melhoria dos resultados.

No Brasil, os números do suicídio não são tão altos se comparados aos de outros países, sendo fácil constatar países desenvolvidos, ou mesmo emergentes, com taxas bem maiores (71ª posição no mundo, taxa bem mais baixa do que França, Alemanha, Reino Unido, Itália, Suíça, Bélgica, Áustria, Canadá, Estados Unidos e Japão) (SANTOS; BARBOSA, 2017). Isso não exclui o fato de as taxas serem crescentes e cada vez mais custosas aos sistemas públicos de saúde quando não passam de tentativas ou quando o óbito não é imediato.

Neste sentido, Sgobin (2013) realiza uma análise de custos, de um hospital público, referentes a internamentos decorrentes de tentativas de suicídio com alta intencionalidade e alta letalidade de método. Chega-se à conclusão de que os custos envolvidos nesses tipos de casos podem ser tão altos quanto os da doença com o maior nível de dispêndio econômico no mundo (síndromes coronarianas agudas). A Figura 1 expõe o crescimento anual do número de suicídios no Brasil entre 2000 e 2016, ocorreu uma elevação de 68,6%, vis-à-vis um crescimento populacional de apenas 18,8%. Isso equivale a uma taxa de crescimento média anual de 4%. A taxa de suicídio que era 3,9 casos para cada 100.000 habitantes, em 2000, passou para 5,5 em 2016, um aumento de aproximadamente 42% na proporção de suicídios.

Figura 1 – Evolução do total (eixo da esquerda) e da taxa de suicídio por 100.000 habitantes (eixo da direita) no Brasil de 2000 a 2016



Fonte: Datasus (mortes por suicídio) e IBGE (população).

Ademais, Durkheim (2000), que escreveu acerca do suicídio em 1897, e tantos outros pesquisadores anteriores e posteriores teorizaram e verificaram empiricamente as relações entre os suicídios de áreas específicas e fatores socioeconômicos, com Durkheim (2000) chegando a afirmar que esses fatores são preponderantes sobre todos os outros para explicar o suicídio. Assim, fica evidente a evitabilidade desse tipo de morte mediante políticas públicas.

O suicídio está entre as dez maiores causas de morte no mundo e entre as três maiores para a faixa de 15 a 35 anos, conforme expõem Santos e Barbosa (2017). No caso do Brasil, Gonçalves et

al. (2011) observam que as taxas de suicídio são crescentes e a quantidade ainda incipiente de trabalhos de pesquisa envolvendo o assunto não exaurem lacunas nesse entendimento, como as relacionadas ao efeito contágio do suicídio. Este último, segundo Durkheim (2000), é conceituado como a forma com a qual o ato de suicídio pode influenciar outros em situações psicológicas e socioeconômicas parecidas, geralmente próximas geograficamente, devendo ser explorada por técnicas de análise espacial. Gonçalves et al. (2011) afirmam que, não apenas na literatura nacional, mas também na internacional, o uso da econometria espacial para abordar esse tema é uma lacuna a ser preenchida.

Dessa forma, o objetivo desta pesquisa é explorar no Rio Grande do Norte os determinantes do suicídio no âmbito do que é enfatizado por Durkheim (2000) e Gonçalves et al. (2011). Apesar das maiores taxas de suicídio serem observadas na região Sul e Centro-Oeste, nos estados do Nordeste elas estão crescendo e, como mostram Santos e Barbosa (2017), dos dez municípios com maiores taxas de suicídio do Nordeste, quatro são do RN, sendo que três desses estão entre as cinco maiores taxas, e a maior é São José do Seridó-RN. O próprio RN é o terceiro estado em taxas de suicídio da região, cabendo, portanto, um estudo que explore mais profundamente suas causas e que forneça uma aplicação do ferramental econométrico-espacial.

Além desta introdução, a seguir é realizada uma revisão da literatura, inicialmente delineando aspectos teóricos e em seguida ressaltando outras análises empíricas na área. Na terceira seção, é explorada a metodologia da verificação empírica que se seguirá, revelando o ferramental utilizado. Na quarta seção, são expostas características referentes à base de dados trabalhada, bem como algumas análises descritivas iniciais. A quinta seção está destinada à exposição e interpretação dos resultados da pesquisa, que está dividida em, inicialmente, Análise Exploratória de Dados Espaciais (AEDE) e, por último, na análise do modelo econométrico-espacial. Ainda, são realizadas considerações finais acerca da metodologia, do modelo estimado e de sugestões derivadas dos resultados, em conclusão, na seção final.

2. Revisão da literatura

Nas subseções seguintes, será feita uma breve revisão da literatura sobre o suicídio. Na primeira, questões teóricas iniciais acerca do tipo de estudo serão delimitadas e, em seguida, as relações deste tipo serão enfatizadas com formulações teóricas compreendidas para o suicídio. Na segunda subseção, são exploradas investigações empíricas realizadas análogas a essa pesquisa, destacando aplicações nacionais e metodologicamente semelhantes.

2.1. Revisão teórica

Um primeiro ponto a ser levantado deve ser a definição do principal objeto deste estudo, o suicídio. Uma conceituação formal e rigorosa torna-se necessária à ciência para que não se incorra no risco de cair em dubiedades e má identificação do que é ou não enquadrado no objeto de estudo; tal como pode acontecer caso seja tomado apenas em seu sentido popular subjetivo, e não delineado e categorizado. Assim, Durkheim (2000) investiga as controvérsias que poderiam surgir dessa definição e chega a seguinte formulação: “Chama-se suicídio todo o caso de morte que resulta direta ou indiretamente de um ato, positivo ou negativo, realizado pela própria vítima, e que ela sabia que produziria esse resultado” (DURKHEIM, 2000, p. 14). Para além disso, no momento da análise empírica desta pesquisa, deve ser tomado um conceito ainda mais objetivo e pragmático, característico dos próprios dados, que caracteriza suicídio como a morte decorrente de lesão autoprovocada intencionalmente, cuja categorização é realizada a partir da Classificação Estatística Internacional de Doenças e Problemas Relacionados à Saúde (CID-10, X60-84).

Além disso, Durkheim (2000) identifica a possibilidade da ocorrência de suicídios em padrões definidos no espaço. O termo “espaço” tem um caráter problemático e complexo para se definir e para ser alcançada uma convenção ou um senso comum em torno do seu significado. Ao se observar as diversas vertentes de abordagem do espaço na ciência regional, como enfatizam Cruz et

al. (2011), é perceptível a adoção, por alguns, de uma ideia de espaço mais palpável ligado à matemática euclidiana e/ou às coordenadas geográficas, de forma mais aplicável às formulações de modelos teóricos e empíricos, enquanto outros tentam lançar mão de uma ideia de espaço mais abrangente e subjetiva que resulta de uma interação de diversos elementos.

Contudo, não será tema deste trabalho discutir questões elementares como a definição de espaço, já que, como afirmam Cruz et al. (2011), o conceito a ser usado depende do enfoque aplicado sobre o objeto de análise. Portanto, será adotada a definição de “espaços territorialmente contíguos inseridos em espaços nacionais sobre os quais há possibilidades concretas de intervenção e de levantamento de informações individualizadas” (CRUZ et al., 2011, p. 45). Tal definição será adotada, sobretudo, tendo em vista que carrega uma maior capacidade de aplicabilidade de políticas e de mensuração de seus efeitos com objetividade.

Praticamente todos os dados passíveis de estudo e referentes a algum fenômeno têm uma localização espacial, desde os dados com elementos sociais e econômicos mais comuns aos aqui estudados, em um espaço bidimensional no âmbito da geografia física, até os estudos em topografias mais abstratas e complexas, como dados de sinapses no espaço tridimensional do cérebro (ASSUNÇÃO, 2001). Seja como for, fica evidente a necessidade de constatação da localização espacial dos objetos tratados em estudo.

Entretanto, a mera característica dos dados provirem de uma localização específica não faz das informações estatísticas extraídas uma técnica de estatística espacial, assim como também dados associados ao tempo não necessariamente serão uma técnica de séries temporais, como enfatiza Assunção (2001). Para tal ser conceituado, as coordenadas espaciais devem estar no processo de coleta dos dados e também em sua descrição e análise. Assim, segundo Monteiro et al. (2004), a análise espacial deve realizar a verificação de propriedades e das características de relações do objeto de estudo, considerando de forma explícita a localização dos fenômenos estudados, incorporando o espaço no modelo utilizado.

A análise espacial pode ser definida como um “conjunto de procedimentos encadeados cuja finalidade é a escolha de um modelo inferencial que considere explicitamente os relacionamentos espaciais presentes no fenômeno” (MONTEIRO et al., 2004, p. N/A). Para que a espacialidade dos dados seja considerada em um modelo, de modo a torná-lo uma análise estatística espacial, deve ser verificado se suas características espaciais afetam as variáveis em observação, ou seja, se existe algum tipo de associação ou relação espacial significativa nos objetos de estudo.

À vista disso, no que concerne à variável suicídio, Émile Durkheim foi quem compôs uma das primeiras e mais importantes sistematizações teóricas realizadas a respeito, com o livro *Le Suicide*, publicado em 1897, no qual o autor busca explorar os fatores individuais e, principalmente, sociais que determinam o suicídio. Em Durkheim (2000), o fator espacial e o suicídio se relacionam por meio da imitação. Esse é um conceito que tem natureza psicológica e se revela quando um ato realizado sucede outro semelhante anteriormente representado, “sem que entre essa representação e a execução se intercale nenhuma operação intelectual, explícita ou implícita, sobre as características do ato reproduzido” (DURKHEIM, 2000, p.138). Assim, é enfatizado, portanto, o caráter irracional da motivação do suicídio pela imitação, em uma espécie de contágio que parece ficar mais evidente quando há proximidade espacial entre os atos.

Por mais que a imitação, como Durkheim (2000) afirma, por si só “raramente seja uma explicação suficiente” para o ato suicida, por isso mesmo chega-se à conclusão que ela não é um fator original, mas sim um fator de reforço a outros motivos. Logo, se essa relação existe, as ocorrências iniciais devem ter um centro de irradiação que influencia a propagação para regiões próximas, sendo sua distribuição perceptível geograficamente. Dessa forma, advoga-se que essa análise deve ser feita a partir do uso de mapas e métodos que o “interroguem” de maneira correta. Assim, ele o faz como exemplo, para a investigação do caso dos distritos franceses, e também para a Europa, apesar de ainda não gozar de ferramentas apropriadas, como as computacionais, e dos métodos mais avançados e eficazes como são os atuais.

Além das relações espaciais do suicídio, nesta pesquisa pretende-se verificar a hipótese de relação dessa variável com o nível de urbanização e desenvolvimento e, ademais, com outras variáveis socioeconômicas. A principal análise realizada por Durkheim (2000) é acerca dos fatores que vão adiante daqueles individuais que são considerados do campo da psicologia. Busca-se, portanto, explorar elementos que influenciam a realização do suicídio de uma forma genérica, conjunta, como característica de uma sociedade. A parcela de suicídios cometidos preponderantemente por causa destes fatores é o que passa a se chamar de taxa social do suicídio.

Assim, o autor numera variados tipos de causas sociais para o suicídio e constrói uma análise que evidencia os mecanismos pelos quais funcionam cada um desses fatores. A sistematização de Durkheim (2000) é realizada de tal forma que classifica os tipos de suicídio relacionando-os com suas causas sociais, nomeadamente: o suicídio egoísta, que está relacionado ao maior nível de independência, exclusão e isolamento social; o altruísta, que ressalta a integração e espírito de corpo social; e o anômico, relativo a questões patrimoniais, conjugais e a fatos específicos.

Mais especificamente, Durkheim (2000) observa que o aumento da independência pessoal de um indivíduo é um fator influenciador para cometer o suicídio e que um nível financeiro mais elevado de renda e um alto grau de educação são elementos que contribuem para essa autonomia. Por outro lado, uma situação mais crônica de carência financeira e educacional podem tornar os indivíduos mais resignados. Dessa forma, o nível de desenvolvimento de uma sociedade parece resumir de forma adequada fatores que levam a uma maior independência dos indivíduos no geral. Outro caso é a relação com a urbanização, nesse caso, afirma-se que, em zonas urbanas, aparentemente, ocorrem mais casos do que em zonas rurais, inevitavelmente ainda apresentando uma relação espacial, abreviada pela taxa de urbanização. Exploram-se ainda vários outros fatores: avanço da idade, sexo masculino e fatos específicos como guerras, fenômenos naturais e crises econômicas que têm uma relação positiva com o suicídio; elementos que levam à integração social e ao fortalecimento de laços familiares, e religiosidade, sobretudo quando essa é relacionada a um maior controle moral, apresentam relações inversas.

2.2. Revisão empírica

Podem se encontrados poucos trabalhos que buscam verificar a imitação como um determinante do suicídio, sobretudo observando o Brasil, apesar da existência de alguns que buscam verificar a relação com outras variáveis. Como já foi visto em Durkheim (2000), a verificação do contágio deve ser feita via observação de mapas e com métodos apropriados para sua análise. Atualmente, o método mais apropriado para isso é, certamente, a Análise Exploratória de Dados Espaciais (AEDE) e a econometria espacial.

Neste sentido, Gonçalves et al. (2011) realizam uma profusa análise espacial do suicídio para o Brasil até o ano 2000, evidenciando o efeito contágio nas microrregiões brasileiras a partir da metodologia de AEDE e econometria espacial. Verifica-se um índice de autocorrelação espacial, *I* de Moran, para o suicídio sempre positivo e significativo que está em torno de 0,52 e 0,57. Além de perceber, a partir de indicadores locais (LISA), a presença de *clusters* com padrão de altas taxas de suicídio no sul e parte do centro-oeste. Na modelagem econométrica, para as variáveis explicativas, fez-se o uso de uma análise fatorial que condensou 17 variáveis em 4 fatores, que compreendem os temas pobreza, grau de ruralização, ensino superior e mulher. Seus resultados sugerem que a defasagem espacial significativa confirma o efeito contágio, a pobreza tem uma relação inversa com o suicídio, o grau de ruralização tem um efeito direto e as variáveis relativas ao sexo e à educação não parecem ter relevância na explicação.

Em outro estudo, também para todo o Brasil, mas agora considerando os municípios como unidades espaciais e o ano de 2010 como unidade de tempo, Dantas et al. (2018), a partir do uso da AEDE, examinam o efeito contágio por medidas de associação espacial e encontram um *I* de Moran que equivale à, aproximadamente, metade do encontrado por Gonçalves et al. (2011). Além disso, a partir de um *I* de Moran global e índice LISA bivariados, e mais regressões lineares bivariadas, verificou-se a relação entre o suicídio e 9 variáveis socioeconômicas apanhadas do Atlas de

Desenvolvimento Humano da PNUD, tais como IDHM, vulnerabilidade, razão de dependência e desemprego, concluindo que o suicídio tem uma fraca autocorrelação espacial, bem como uma pouco observada correlação espacial, e não espacial, com as variáveis socioeconômicas. Metodologicamente semelhante, ao explorar os conglomerados de suicídio, Santos e Barbosa (2017) fazem o mesmo para o Nordeste brasileiro, com exceção das regressões lineares bivariadas, e obtêm resultados parecidos, concluindo que o suicídio tem uma distribuição espacial aleatória e não tem relação espacial com variáveis socioeconômicas.

Não foram observados outros trabalhos que usam metodologias de análise espacial, apesar do efeito contágio ter sido observado também por Cutler et al. (2001), o qual percebe que esse efeito é pouco evidente em adultos e idosos, sendo uma provável justificativa para menor importância das relações sociais para esse grupo. Ao contrário dos suicídios ocorridos e tentados pelos mais jovens que, segundo casos específicos investigados pelo autor, o contágio é mais evidente.

Loureiro et al. (2010) não analisam o efeito contágio, mas buscam evidenciar a relação do suicídio com fatores econômicos. Eles observam que as taxas de suicídio têm uma maior relação com idades mais baixas e, ainda mais, variáveis econômicas têm maior impacto sobre o suicídio em indivíduos mais jovens. Esse resultado é encontrado com uma metodologia de dados em painel para os estados brasileiros, de 1981 a 2006, utilizada para verificar a influência de variáveis econômicas sobre o suicídio. Seus achados concordam com a tese verificada de que a felicidade está atrelada ao consumo relativo e não ao consumo absoluto, ao observar que a desigualdade tem uma relação direta e a pobreza tem uma relação inversa.

Com dados *cross-section*, Shikida et al. (2007) também avaliam o suicídio nos estados brasileiros, contudo partindo da teoria econômica do suicídio e estimando um modelo baseado em formulações teóricas estritamente estabelecidas. Eles concluem que a evidência empírica tem o ajuste esperado ao modelo teórico, verificando uma relação cúbica com a idade, *dummies* para a região sul e centro-oeste apresentaram grande relevância, bem como os gastos com saúde. Por outro lado, Schnitman et al. (2011), com uma análise de correlação e regressão clássica, não encontram muitas evidências de relações com fatores socioeconômicos e concluem que fatores biopsicológicos devem ser preponderantes na motivação do suicídio.

Em uma análise para os municípios do Brasil e do Rio Grande do Sul, especificamente este estado por ter os municípios com maiores taxas de suicídio, para o ano 2010, Fraga et al. (2016) utilizam ferramentas, como a Regressão de Poisson, Binomial e Quantílica, para verificar a influência de variáveis socioeconômicas sobre o suicídio. Foram observadas relações diretas com a pobreza, a cor/raça branca e indígena e, destacadamente, o nível de dependência agropecuária. Este último o autor associa às transformações recentes ocorridas na área rural dos municípios brasileiros e às questões econômicas, como a pobreza, a mudança nas formas de produção (máquinas e produtos químicos), o endividamento, a concentração da terra, o êxodo rural e o isolamento desse estrato social.

Dutra e Roehe (2013) realizaram pesquisa qualitativa sobre os cinco municípios com maiores taxas de suicídios entre agricultores do Rio Grande do Norte, a partir da aplicação de questionários a pessoas que vivem da agricultura nesses locais e são parentes de vítimas de suicídio. O produto de suas entrevistas comprova parte das conclusões tiradas por Fraga et al. (2016) acerca da condição do indivíduo no meio rural, o qual passa por situações degradantes do ponto de vista social e econômico. Para além disso, Dutra e Roehe (2013) interpretam as informações à luz da filosofia e psicologia, em uma abordagem existencial.

Por fim, interessa como recomendação para o aprofundamento empírico acerca do suicídio, sobretudo considerando aplicações em diversas partes do mundo, o artigo de revisão de Rehkopf e Buka (2006), que faz uma vasta análise da literatura empírica existente até então, acerca da associação do suicídio com características socioeconômicas de áreas geográficas. A revisão é feita tomando 221 análises encontradas, consideradas de alta qualidade, estratificando por características, como o tamanho da área agregada, as medidas socioeconômicas, a região de estudo e o desenho da modelagem e, por conseguinte, analisando o percentual de tipo de resultado, como positivo,

negativo e nulo. Como conclusão, há a confirmação da grande heterogeneidade de associações encontradas, conseqüente dos discrepantes resultados encontrados em diversos estudos, apesar de que tendências puderam ser observadas em conformidade com as estratificações usadas.

3. Metodologia

Monteiro et al. (2004) observam que, geralmente, a modelagem é precedida do que é comumente chamado de análise exploratória de dados espaciais (AEDE), a qual busca apresentar visualmente em gráficos e mapas os padrões espaciais e as relações de dependência. Neste trabalho, além de estatísticas descritivas e *box plots*, uma AEDE desenvolvida fará uso de diagramas de dispersão (*Moran scatterplot*), indicadores de autocorrelação espacial e análise de *clusters* espaciais. Os dois procedimentos mencionados inicialmente ainda não consideram os fatores espaciais, mas são capazes de fornecer uma ideia inicial de características importantes das observações. Seguindo Anselin (2005), os últimos são formas gráficas de observar medidas de tendência central e distribuição espacial dos dados e verificar a existência de *outliers* e pontos de alavancagem¹.

O *Moran scatterplot*, mais especificamente, é a primeira ferramenta utilizada, nesta pesquisa, que irá considerar o fator espaço, já que plota em um plano cartesiano a dispersão dos desvios da variável analisada contra sua defasagem espacial. Nos seus quadrantes, é possível identificar o padrão de associação em quatro categorias, alto-alto (AA), baixo-baixo (BB), alto-baixo (AB) e baixo-alto (BA), que indicam observações que têm valores altos (ou baixos) rodeados por valores baixos (ou altos) e vice-versa. Além de existirem *outliers* (nos quadrantes AB e BA) e pontos de alavancagem (nos quadrantes AA e BB), tipicamente categorizadas quando observações apresentam valores maiores que dois desvios-padrão (ANSELIN, 2005).

Anteriormente a isso, deve ser identificada a matriz de ponderação espacial W de dimensão $n \times n$, definida por Anselin e Rey (2014), como o componente mais importante da análise *cross-section* de dependência espacial, pois expressa a estrutura de vizinhança entre as observações. Apesar das diferentes formas possíveis de obter padrões de vizinhança, a construção de W realizada será baseada na relação binária expressa no padrão *queen* de primeira ordem, que considera vizinhas de uma unidade espacial todas as outras que compartilham com a primeira a mesma borda ou vértice, ou seja, regiões que fazem qualquer tipo de fronteira entre si. Assim, cada termo de peso espacial w_{ij} , da matriz W terá valor 1 sempre que i for vizinho de j e 0 caso contrário. Por fim, a matriz W considerada será a normalizada, em que o valor de cada w_{ij} será dividido pelo somatório de sua própria linha. Há ainda a possibilidade do uso de matrizes de ponderação que identificam padrões diferentes de vizinhança, tais como a matriz *root*, que considera como vizinhas apenas aquelas regiões que dividem o mesmo vértice de fronteira, as matrizes de distância, que avaliam como vizinhos aqueles mais próximos dentro de uma distância definida arbitrariamente ou, também, as matrizes de k -vizinhos, que consideram apenas como regiões vizinhas as k mais próximas, sendo k também arbitrário.

Após a obtenção da matriz de pesos espaciais, ainda segundo Anselin e Rey (2014), é possível obter o vetor das defasagens espaciais da variável objetivo, y , com o somatório do produto de suas observações pelos pesos espaciais, w_{ij} , para cada unidade espacial. Em suma, será obtido através de Wy e isso irá descrever, portanto, uma soma ou média ponderada dos valores vizinhos para essa variável. Isto posto, torna-se possível realizar todas as verificações necessárias neste trabalho que consideram o fator espaço.

¹ Torna-se necessário esse tipo de análise, dada a relevante influência que observações discrepantes podem exercer sobre o modelo, deturpando estimativas e testes. Os *outliers* espaciais são observações que destoam do resto em magnitude e demonstram relação inversa entre as vizinhanças, não seguindo o mesmo padrão de dependência espacial. São definidas, portanto, nas extremidades dos quadrantes de padrão alto-baixo e baixo-alto. Já os pontos de alavancagem seguem o mesmo padrão de associação espacial, mas exercem influência exacerbada sobre a determinação da medida de correlação espacial da amostra, por isso, são identificados nas extremidades dos quadrantes alto-alto e baixo-baixo.

Feito isso, será possível obter o nível de associação espacial dos dados. Como medida global, será considerada a estatística I de Moran, que mede o grau de autocorrelação da variável no espaço. Seu valor pode variar entre -1 e 1, quanto maior, mais positivamente autocorrelacionada será a variável; caso negativo, a variável terá relação inversa com suas observações vizinhas e valores próximos a zero indicarão pouca associação espacial, ou aleatoriedade. Seguindo Anselin (1995), e apenas transformando sua notação para a abordagem vetorial, o valor de I deve ser descrito pela Equação 1.

$$I = \frac{n}{S_0} \frac{z'Wz}{z'z} \quad (1)$$

Nessa equação, o número de observações, ou unidades espaciais, é dado por n . O valor de S_0 representa a soma de todas linhas e colunas de W , cujo resultado, neste caso, é um escalar igual a n , já que a matriz de pesos espaciais é normalizada nas linhas. O vetor z de dimensão n é identificado como os desvios da variável y em relação à média ($y_i - \bar{y}$). A validação estatística é baseada em um teste de pseudo-significância, que a partir de permutações aleatórias recalcula o índice I muitas vezes e gera uma distribuição de referência, dessa forma, o valor calculado pode ser comparado com a distribuição e computada uma pseudo-significância (ANSELIN, 2005).

Contudo, o índice global expõe apenas uma única medida para o conjunto de dados como um todo. Cabe, portanto, ressaltar a existência de padrões de associação local, que podem ter um padrão espacial estatisticamente significativo e diferente do exposto pelo indicador global. Para tanto, Anselin (1995) apresenta os indicadores LISA (*Local Indicator of Spatial Association*) que, a partir da decomposição dos indicadores globais, I de Moran neste caso, são capazes de fornecer informações acerca de *clusters* espaciais e seus padrões, fornecendo uma espécie de indicador para cada unidade espacial, conforme descrito na Equação 2.

$$I_i = z_i \sum_j w_{ij} z_j \quad (2)$$

Para obter um indicador de associação espacial entre duas variáveis, deve ser considerada a relação dos valores da segunda variável nos vizinhos com a primeira variável na unidade espacial em questão. Anselin et al. (2002) demonstram que para obter isso basta substituir na Equação 1 a defasagem espacial dos desvios de y (Wz , agora com z identificado pelos subscritos das duas variáveis) pela defasagem dos desvios da nova variável x (Wz_x), conforme a Equação 4. De maneira análoga, isso pode ser feito para os indicadores locais (Equação 4).

$$I_{yx} = \frac{n}{S_0} \frac{z'_y W z_x}{z'_y z_y} \quad (3)$$

$$I_{yx}^i = z_y^i \sum_j w_{ij} z_x^j \quad (4)$$

Dada a dificuldade de observar os resultados dos indicadores de associação local, sua interpretação fica facilitada com a plotagem do indicador de cada unidade espacial em um mapa. Portanto, a análise desses resultados é realizada, seguindo Anselin (2005), com o mapa de *clusters*, que identifica em cada unidade espacial o padrão de vizinhança, classificada conforme o observado em um diagrama de dispersão, com os quadrantes AA, BB, AB e BA, desde que tenha significância estatística. Ademais, também pode ser observado um mapa de significância que expõe o padrão de pseudoníveis de significância estatística gerado a partir de permutações, de forma análoga ao realizado ao I de Moran global.

A fim de realizar a modelagem para a variável estudada e verificar a relação entre essa e outras variáveis socioeconômicas, recorre-se ao modelo de regressão para dados *cross-section*, no caso, um modelo clássico estimado pelo método dos mínimos quadrados ordinários (MQO), representado como na Equação 5, na qual y é explicado por X , que representa uma matriz de covariáveis associada a um vetor de parâmetros β e mais um vetor de componentes aleatórios ε , normalmente distribuídos com média zero e variância constante $\sigma^2 I$.

$$y = X\beta + \varepsilon, \quad \text{onde: } \varepsilon \sim N(0, \sigma^2 I) \quad (5)$$

Como expõe LeSage e Pace (2009), para que esse modelo seja válido, é necessária a pressuposição observada sobre ε , a qual implica que todos os termos ε_i para cada unidade espacial i sejam independentes dos demais ($E(\varepsilon_i \varepsilon_j) = 0$) e não correlacionados com X ($E(\varepsilon|X) = 0$). No entanto, torna-se forçoso considerar essa simplificação caso o processo gerador dos dados analisados reflitam um contexto de dependência espacial, que poderia fazer com que os estimadores tivessem problemas de viés, inconsistência e/ou ineficiência. Consequentemente, cabe uma investigação sobre os erros da regressão para verificar a existência de processos desse tipo.

Para tal fim, pode ser aplicado um teste de Moran sobre os resíduos para verificar a existência de autocorrelação espacial. Se o I de Moran se mostrar estatisticamente significativo, deve ser definido um novo modelo de estimação que considere o processo de geração específico dos dados. O modelo a ser definido neste trabalho faz parte do que Monteiro et al. (2004) classificam como aqueles que “tratam a estrutura espacial de forma global”, isto é, com um único parâmetro, sobretudo SAR e SEM, e não com parâmetros que variam continuamente no espaço. Nas Equações 6, 7 e 8, podem ser observadas suas especificações.

$$y = \rho W y + X\beta + \varepsilon \quad (6)$$

Em um modelo de defasagem espacial autorregressivo (*Spatial Autoregressive Model - SAR*) observado na Equação 6, como é exposto por LeSage e Pace (2009), é modelado o efeito de dependência espacial inerente à variável dependente, portanto, é adicionado como variável explicativa o termo de defasagem espacial de y , representado pelo vetor $W y$, cujo efeito sobre a variável dependente é materializado pelo coeficiente ρ .

$$y = X\beta + \mu, \quad \text{onde: } \mu = \lambda W \mu + \varepsilon, \quad \text{logo:} \quad (7)$$

$$y = X\beta + (I_n - \lambda W)^{-1} \varepsilon$$

No modelo de erro espacial (*Spatial Error Model - SEM*), tenta-se modelar efeitos espaciais não incorporados ao modelo, como os de variáveis autocorrelacionadas espacialmente não consideradas, de forma que o efeito se apresenta como uma perturbação no termo de erro. Portanto, na Equação 7, na qual I_n é uma matriz identidade, é possível observar que o termo de erro μ , além de conter o erro aleatório ε , apresenta um termo autorregressivo μ cujo efeito é captado pelo coeficiente λ , que passa a ser modelado. Os processos observados nos modelos SAR e SEM também podem aparecer juntos e serem considerados em um único modelo (*spatial autoregressive confused - SAC* ou o *spatial auto-regressive and moving average - SARMA*), especificado na Equação 8.

$$y = \rho W y + X\beta + (I - \lambda W)^{-1} \varepsilon \quad (8)$$

Para definição do modelo espacial, será seguida a abordagem clássica do procedimento proposto por Florax et al. (2003) que consideram as seguintes etapas como critério de seleção: (1) estima-se o modelo clássico por MQO; (2) testa-se a hipótese de dependência espacial a partir dos

testes de Multiplicador de Lagrange (LM), descritos na Equação 9², na qual LM_ρ designa a estatística que tem como hipótese alternativa o uso da defasagem espacial de y no modelo, e LM_λ , que aponta o uso do modelo de erro espacial caso a hipótese nula seja rejeitada; (3) caso ambos os testes não sejam significantes, estima-se o modelo clássico por MQO; (4) senão, caso ambos sejam estatisticamente significantes, estima-se a especificação do que for apontado como mais significativo; (5) por outro lado, se apenas LM_ρ for significativa, estima-se o modelo SAR; (6) caso contrário, se nem isso acontecer, estima-se o modelo SEM. Além disso, se for conveniente, pode ser utilizada uma estratégia de especificação híbrida que incorpora as formas robustas do teste LM , sobretudo como critério de desempate na etapa 4.

$$LM_\rho = \frac{(\hat{\varepsilon}'Wy/\hat{\sigma}^2)^2}{nJ}, \quad LM_\lambda = \frac{(\hat{\varepsilon}'W\hat{\varepsilon}/\hat{\sigma}^2)^2}{T} \quad (9)$$

Como já foi afirmado, a estimação de modelos espaciais por MQO pode levar a estimadores viesados, inconsistentes e/ou ineficientes, devido a problemas de endogeneidade envolvendo as variáveis espaciais, Wy e $W\mu$. Como alternativa, a literatura largamente sugere o uso do método da máxima verossimilhança (MV), que é capaz de gerar estimadores ρ e λ com propriedades de consistência, eficiência e normalidade assintótica, apesar de também requerer a resguarda de alguns pressupostos, como a normalidade dos resíduos (LESAGE; PACE, 2009).

Após a estimação, a verificação de alguns testes é requerida. Para diagnosticar os resíduos, será aplicado o teste Jarque-Bera, cuja hipótese nula é a normalidade dos resíduos (para o MQO) e os testes Breusch-Pagan (também para MV) e Koenker-Bassett, com hipótese nula de homocedasticidade. Também será possível testar a especificação do modelo espacial, tendo como hipótese alternativa a preferência pelo parâmetro espacial. No caso do uso de variáveis instrumentais (VI), poderá ser verificado o esgotamento dos efeitos espaciais no modelo pelo teste Anselin-Kelejian, cuja hipótese nula é de que não há mais autocorrelação espacial nos resíduos do modelo (ANSELIN, 2005).

4. Base de dados

Para a análise realizada, foram utilizados dados *cross-section* colhidos da base de dados socioeconômicos do Atlas de Desenvolvimento Humano do Programa das Nações Unidas para o Desenvolvimento (PNUD), que faz uso do Censo Demográfico de 2010. Portanto, todas as variáveis extraídas dessa base serão referentes a esse período e serão utilizadas como covariáveis. Os dados de população relacionados à taxa de suicídio são também dessa fonte. Porém, os dados acerca do número de suicídios são extraídos da base do Datasus, que tem como fonte o Sistema de Informações de Mortalidade (SIM), referente aos anos de 2006 a 2015.

As outras variáveis utilizadas na análise, presentes no Quadro 1, foram selecionadas buscando captar características socioeconômicas dos municípios que tenham uma possível relação com a taxa de suicídio e em relevante parte tendo como embasamento a literatura teórica e empírica no processo de seleção.

Assim, como motivação para utilização, variáveis como o índice de desenvolvimento e de Gini, a renda e a proporção da população com ensino superior completo, que são vistos como elementos que levam à maior independência dos indivíduos, são de largo uso na literatura (GONÇALVES et al., 2011; DANTAS et al., 2018; SANTOS; BARBOSA, 2017; LOUREIRO et al., 2010; SHIKIDA et al., 2007; SCHNITMAN et al., 2011; FRAGA et al., 2016) a urbanização atende a um critério de viés bem mais teórico, como é elucidado em Durkheim (2000); bem como a fecundidade, vista por Durkheim (2000) como um elemento que fortalece os laços familiares, e o decil dos ricos, não

² No qual $\hat{\sigma}^2$ e $\hat{\varepsilon}$ são, respectivamente, a variância e os resíduos estimados, T é o traço da matriz $((W'+W)W)$ e J é um escalar que capta relações entre W , X , $\hat{\sigma}^2$ e os estimadores de MQO.

encontrado em outros trabalhos empíricos, mas que quantifica o quão alto é o nível de riqueza das classes mais abastadas enfatizadas por Durkheim (2000) como mais propensas ao suicídio. A taxa de desemprego (a qual esperava-se ter uma relação inversa com a taxa de suicídio) também é verificada por Dantas et al. (2018), Santos e Barbosa (2017), Loureiro et al. (2010) e Fraga et al. (2016). Ademais, como alternativa ao Gini, a desigualdade é dada pelo índice de Theil (usado por Schnitman et al. (2011)), razão de dependência (usado por Schnitman et al. (2011), Dantas et al. (2018) e Santos e Barbosa (2017)) e pobreza (usado por Schnitman et al. (2011), Loureiro et al. (2010), Gonçalves et al. (2011), Fraga et al. (2016), Dantas et al. (2017) e Santos e Barbosa (2017)).

Quadro 1 – Descrição das variáveis

Variável	Descrição
Taxa de Suicídio	Taxa da média anual de óbitos ocorridos entre 2006 e 2015 para cada 100.000 habitantes (classificação do óbito pela CID-10: X60-X84).
Taxa de urbanização	Razão entre a população urbana e a população total.
IDH-M	Índice de Desenvolvimento Humano Municipal.
Renda <i>per capita</i>	Renda <i>per capita</i> média de residentes em domicílios particulares permanentes.
Taxa de desocupação	Taxa de desocupação da população de 18 anos ou mais de idade.
Taxa de fecundidade	Número médio de filhos que uma mulher deverá ter ao terminar o período reprodutivo.
Índice de Gini	Grau de desigualdade na distribuição de indivíduos segundo a renda <i>per capita</i> .
Ensino superior	Proporção da população de 25 anos ou mais com ensino superior completo.
Índice de Theil	Desigualdade na distribuição de indivíduos segundo a renda domiciliar <i>per capita</i> .
Razão de dependência	Razão entre os <15 e >65 anos em relação aos >15 e <64 anos.
Pobreza	% vulneráveis à pobreza (menos de 1/2 salário mínimo de renda domiciliar <i>per capita</i>).
Decil dos ricos	Renda <i>per capita</i> mínima do décimo mais rico.

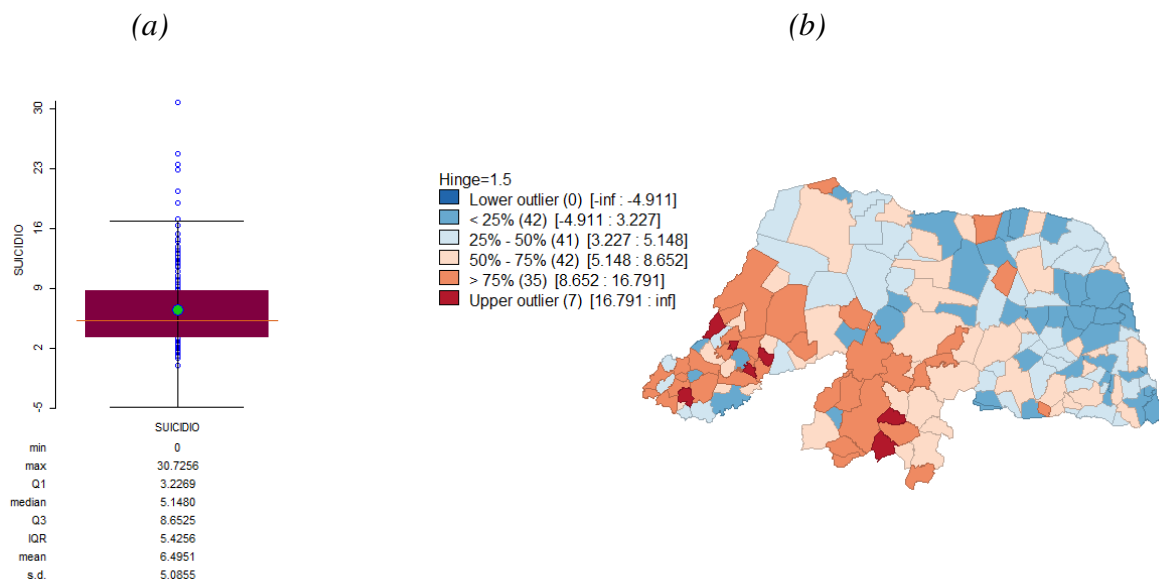
Fonte: Elaboração própria, com base nas descrições da PNUD e Datasus.

No *box plot* da taxa de suicídio, apresentado na Figura 2, é possível notar visualmente o grau de dispersão da variável. A caixa, que contém 50% das observações mais prováveis, ou seja, 83 municípios, está delimitada pelas taxas de suicídio nos valores de 3,22 (1º quartil - Q1) e 8,65 (3º quartil - Q3). A linha laranja representa o município de Serra Negra do Norte, que indica o valor mediano 5,14, ou seja, que divide a amostra da variável ao meio. A média observada para os municípios é de 6,49 suicídios para cada 100.000. O valor mínimo foi observado em 11 municípios, que não relataram nenhum caso de suicídio, e o valor máximo foi observado no município de São José do Seridó, com uma taxa de 30,72.

Além disso, acima da linha superior do *box plot*, podem ser constatadas 7 observações discrepantes caracterizadas como *outliers* altos, nomeadamente: José da Penha (17,04), Frutuoso Gomes (18,90), Rodolfo Fernandes (20,37), Rafael Godeiro (22,85), Ouro Branco (23,41), Viçosa (24,72) e São José do Seridó (30,72). O último foi verificado também por Santos e Barbosa (2017) como a maior taxa de suicídio do Nordeste. Esses municípios podem ter suas localizações identificadas pelos municípios de cor marrom escura do *box map* na Figura 2, que apresenta informações semelhantes ao *box plot*, mas localizadas geograficamente.

Quanto à espacialidade das informações, os dados caracterizam cada um dos municípios do Rio Grande do Norte como observações e unidades espaciais, se tratando, portanto, de um estudo ecológico, em que a unidade de análise é uma população ou grupo de indivíduos, e não o indivíduo propriamente dito. Para esse tipo, Gonçalves et al. (2011) mencionam a possibilidade de se incorrer na “falácia ecológica”, na qual a generalização pode ser indevida, eles alertam que não necessariamente as relações refletirão a situação dos indivíduos especificamente. Contudo, além de esses serem os dados disponíveis, são também utilizados para formulação de políticas públicas na área da saúde, embasando seu potencial de relevância.

Figura 2 – Plotagem e distribuição da taxa de suicídio dos municípios em quartís: *Box Plot* (a) e *Box Map* (b)



Fonte: Elaboração própria.

A descrição mais específica de cada variável é fornecida no Quadro 1, mais especificamente com relação a taxa de suicídio, o procedimento adotado foi, inicialmente, a realização de uma média anual do valor absoluto de suicídios³ para cada município, ocorridos nos dez anos circundantes a 2010, isto é, entre 2006 e 2015. Assim, a taxa foi obtida pela razão entre esta média e a população municipal referente a 2010 multiplicada por 100.000.

Desta forma, a taxa de suicídio deve ser interpretada como a quantidade média anual de suicídios ocorridos a cada 100.000 habitantes para cada município. Com esse procedimento, de uma média que abarca dez anos, procurou-se amenizar os problemas de grandes flutuações anuais e de eventos raros em populações pequenas mencionados por Fraga et al. (2016), já que poucos municípios possuem uma população maior do que 100.000 habitantes.

Para um melhor esclarecimento das análises dos mapas expostos, serão muitas vezes mencionadas as mesorregiões e microrregiões do Rio Grande do Norte. Assim, no caso das quatro mesorregiões (e respectivas microrregiões em parênteses), essas podem ser identificadas como Oeste Potiguar (Mossoró, Chapada do Apodi, Médio Oeste, Vale do Açu, Serra de São Miguel, Pau dos Ferros, Umarizal), que inclui toda porção oeste do estado até parte do centro e final do norte do litoral; Central Potiguar (Macau, Angicos, Serra de Santana, Seridó Ocidental, Seridó Oriental), toda a região que se inicia ao sul do estado que é ladeado por porções da Paraíba e vai até o lado norte, onde encontra o litoral; Leste Potiguar (Litoral Nordeste, Macaíba, Natal, Litoral Sul), localizado em toda região litorânea leste; e, por fim, Agreste Potiguar (Baixa Verde, Borborema Potiguar, Agreste Potiguar), que fica entre o Central Potiguar e Leste Potiguar, sendo a única região sem contato com o mar.

Observando o *box map*, fica perceptível a predominância de baixos valores nas mesorregiões do Leste Potiguar e do Agreste Potiguar, com essas contendo boa parte dos valores dos dois primeiros quartís, isto é, com a maioria de suas observações com taxas de suicídio abaixo de 5,14. De forma oposta, valores mais altos predominam nas mesorregiões Central Potiguar e Oeste Potiguar, sobretudo nas microrregiões da Serra de São Miguel, Pau dos Ferros, Umarizal, Médio Oeste, Chapada do Apodi, Seridó Ocidental e Oriental e Serra de Santana. Em suma, visualmente, é

³ Óbitos decorrentes de lesões autoprovocadas intencionalmente categorizadas a partir da Classificação Estatística Internacional de Doenças e Problemas Relacionados à Saúde (CID-10, X60-84).

bastante perceptível o aumento das taxas de suicídio espacialmente do litoral em direção ao interior do estado.

Também serão analisados de forma mais específica, na Análise Exploratória de Dados Espaciais (AEDE), o Índice de Desenvolvimento Humano Municipal (IDHM) e a taxa de urbanização, buscando verificar a hipótese de relação entre estas variáveis e a taxa de suicídio.

5. Resultados

Nas próximas subseções, serão explorados os resultados das inferências realizadas.⁴ Em um primeiro momento, serão observadas as medidas de associação espacial, buscando investigar o efeito contágio relacionado aos casos de suicídio nos municípios do Rio Grande do Norte e as relações espaciais deste com os níveis de desenvolvimento e urbanização. Em seguida, proceder-se-á com a modelagem econométrica-espacial, buscando verificar o efeito de variáveis socioeconômicas, conforme a melhor estratégia de estimação for constatada.

5.1. Análise Exploratória de Dados Espaciais

Para uma análise do efeito contágio do suicídio, a primeira estatística a ser analisada é o I de Moran global, pois ele fornece um nível de associação geral entre as unidades espaciais para a variável, considerando a matriz *queen* de ponderação espacial. Dessa forma, foi constatado um valor I igual a 0,2686 para a taxa de suicídio entre os municípios do Rio Grande do Norte, aproximadamente igual ao encontrado por Dantas et al. (2018) e Santos e Barbosa (2017) para os municípios do Brasil e do Nordeste, respectivamente. Uma autocorrelação espacial considerada baixa, contudo, estatisticamente significativa com um pseudo p -valor de 0,001 gerado a partir de 999 permutações em que o I foi recalculado. O valor positivo da autocorrelação espacial, mesmo que baixo, evidencia a existência de um efeito contágio do suicídio entre os municípios, cujos padrões locais devem ser investigados.

Na Figura 3, é observado o Moran *scatterplot*, ou gráfico de dispersão de Moran, em que no eixo X apresentam-se os desvios da taxa de suicídio contra a defasagem espacial desses desvios no eixo Y, conforme a matriz de vizinhança *queen*, escolhida baseando-se nos trabalhos de Gonçalves et al. (2011), Dantas et al. (2018) e Santos e Barbosa (2017). No gráfico central, é possível constatar no primeiro quadrante seis pontos de alavancagem (José da Penha, Frutuoso Gomes, Rafael Godeiro, Ouro Branco, Viçosa e São José do Seridó), que são observações com altas taxas de suicídio discrepantes, rodeadas por uma vizinhança que também tem altos valores, considerando a média. No quarto quadrante, é observado um *outlier* espacial, o município de Rodolfo Fernandes, que tem uma alta taxa de suicídio discrepante, rodeado por municípios com baixas taxas.

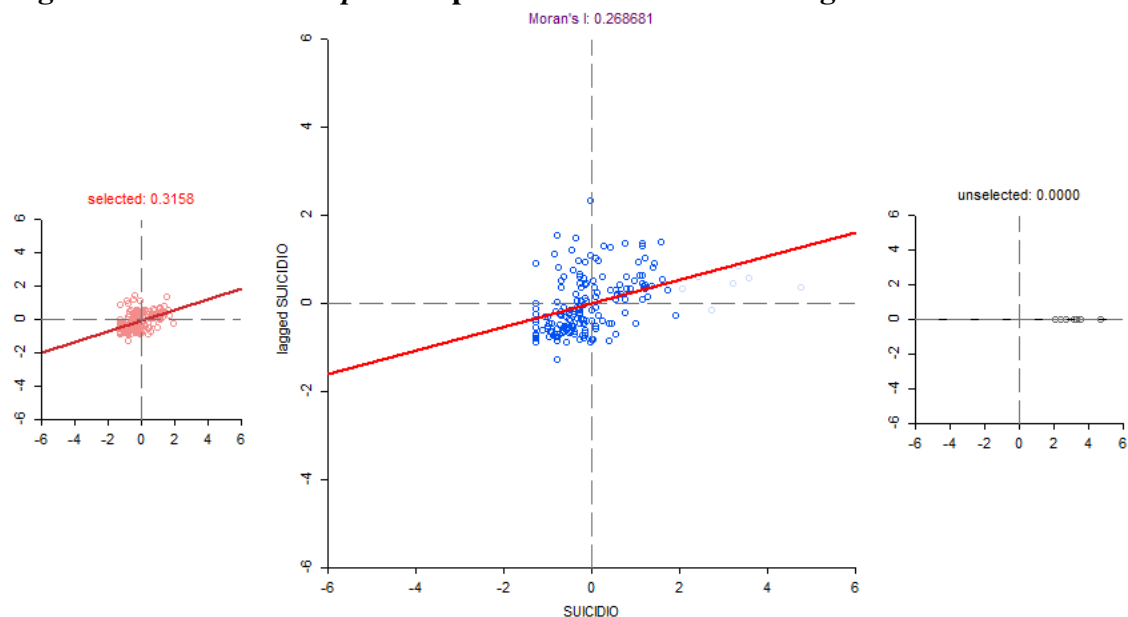
Juntos, esses sete municípios equivalem aos *outliers* globais encontrados no *box plot* e *box map* da Figura 2. Se fossem excluídos da amostra, como no gráfico com pontos vermelhos à esquerda do gráfico central da Figura 3, seria obtido um I de Moran global no valor de 0,3158. A reta vermelha presente nos gráficos deve ser entendida como uma reta de ajuste descrita por $z = \beta Wz + \varepsilon$, no qual $I = \hat{\beta}$, logo o I de Moran de 0,2686 do gráfico central será a inclinação dessa reta de regressão (ANSELIN et al., 2002).

Além disso, também foi avaliado o I de Moran considerando diferentes matrizes de ponderação. Como exemplo, a matriz *rook* rendeu um I significativo de 0,28, bem próximo do gerado usando a matriz *queen*. As matrizes que consideram vizinhos os municípios em raios de 60 km, 100 km e 150 km resultaram em correlações espaciais de, respectivamente, 0,29, 0,22 e 0,17, todos significantes e expondo um resultado óbvio: quanto maior a distância, menor o efeito espacial. Por último, a matriz cujos vizinhos considerados são sempre apenas os dez mais próximos rendeu um I de Moran de 0,36. Contudo, dada a arbitrariedade na especificação dessas outras

⁴ Os resultados da Análise Exploratória de Dados Espaciais (AEDE) foram obtidos usando o *software* GeoDa e os referentes a modelagem econométrica com o GeoDaSpace.

matrizes de ponderação, a matriz *queen*, amplamente já utilizada, será a escolhida para todo o resto das verificações.

Figura 3 – Moran scatterplot: dispersão dos desvios e defasagens da taxa de suicídio



Fonte: Elaboração própria.

A análise de *clusters* da taxa de suicídio confirma, em grande parte, aquilo que já havia sido observado no *box map*. Na Figura 4(a), segundo o indicador local de autocorrelação espacial (LISA), verificam-se *clusters* com padrão de baixas taxas de suicídio com municípios vizinhos semelhantes, identificados em azul escuro, em quase toda região metropolitana de Natal e na maior parte do litoral leste do estado. Esse é o padrão com mais municípios, 28 ao todo.

Outros 9 municípios participam de conglomerados de alta taxa de suicídio nas microrregiões do Seridó, de Pau dos Ferros e de Umarizal. A maioria dos municípios, 117 ao todo, não teve indicadores LISA estatisticamente significantes ao nível de 5%. Assim, pode ser afirmado que altas taxas de suicídio estão mais presentes no interior do estado, enquanto grandes conglomerados de baixas taxas de suicídio se encontram próximos ao litoral, sobretudo no entorno de Natal e da região metropolitana. A maior parte desses *clusters* identificados ressaltam, portanto, o efeito contágio examinado por Durkheim (2000).

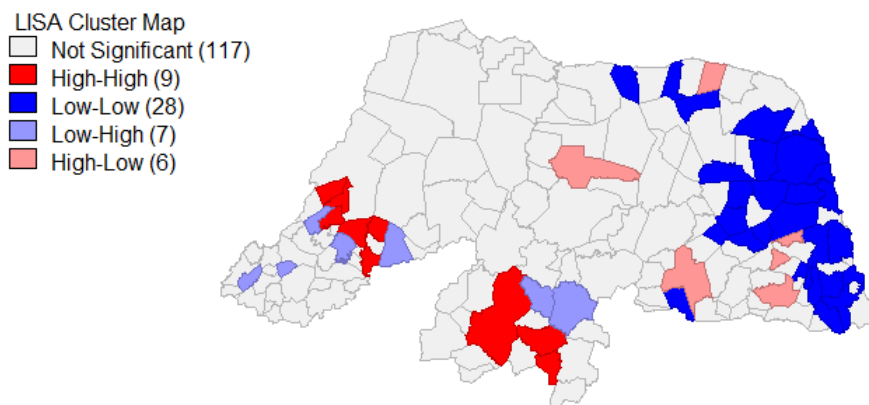
Com intuito de verificar relações espaciais da taxa de suicídio com outras variáveis-chave, foi realizada também uma análise de *clusters* a partir de indicadores LISA bivariados. Na Figura 4(b), é exposto o nível de associação do suicídio com a defasagem espacial do IDHM nos municípios vizinhos. O índice de desenvolvimento humano, usado em praticamente toda a literatura nacional, foi escolhido haja vista que fornece um resumo de informações importantes acerca da renda, longevidade e educação, cuja generalidade pode ser importante para explicar o suicídio não apenas no próprio município.

Assim, foram detectados *clusters* (21 municípios) de baixa taxa de suicídio rodeados por baixos níveis de desenvolvimento municipal nas microrregiões do Litoral Nordeste e da Baixa Verde. Em uma relação semelhante, no Seridó, foram observados municípios (9) onde há altas taxas de suicídio e a vizinhança apresenta altos níveis de desenvolvimento. Examinando a situação de maneira ampla, o *I* de Moran global bivariado calculado para essa relação foi de 0,2022, baixo, mas significativo, ressaltando uma relação direta. Esses resultados podem estar apenas refletindo a relação positiva dos níveis de desenvolvimento com a independência alcançada pelos indivíduos na região conglomerada que, segundo Durkheim (2000), é um fator direto de influência nos atos de suicídio. Apesar de que, foram detectados 12 municípios cujo padrão é de baixas taxas de suicídio

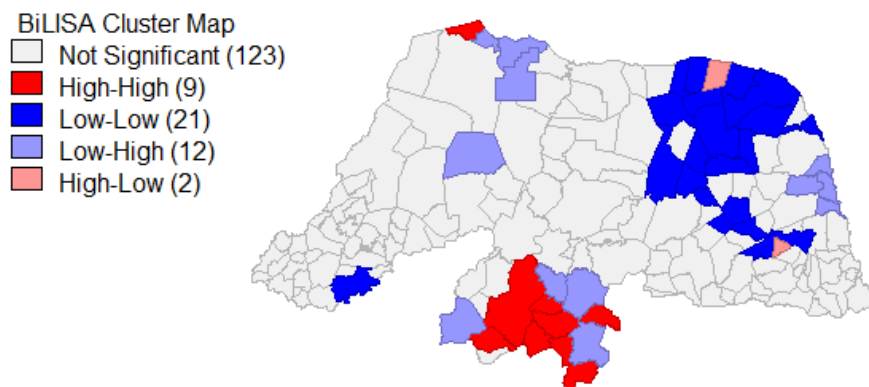
associadas a altos níveis relativos de desenvolvimento. Ao todo, 123 municípios não tiveram indicadores LISA significativos estatisticamente.

Figura 4 – LISA univariado e bivariado: análise de *clusters* da taxa de suicídio, IDHM e taxa de urbanização

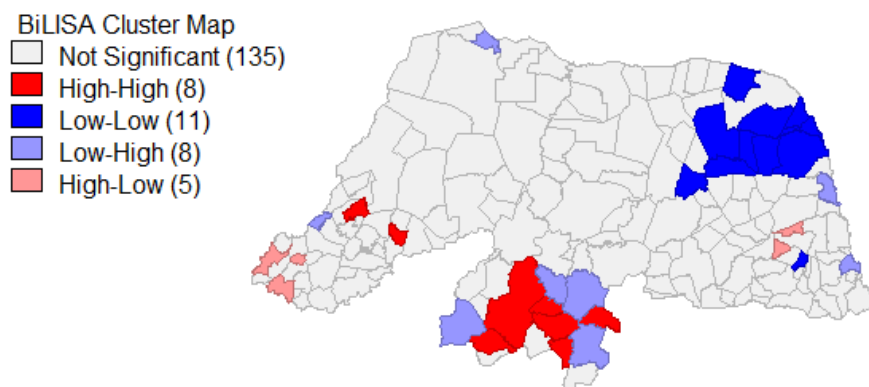
(a) *Taxa de suicídio (univariado)*



(b) *Suicídio x IDHM*



(c) *Suicídio x Urbanização*



Fonte: Elaboração própria.

O mesmo foi realizado para a taxa de urbanização na Figura 4(c), embasando-se na ideia de que esse fenômeno tem um caráter, em grande parte, eminentemente espacial, no qual dificilmente uma cidade se torna altamente urbanizada sem que suas vizinhas sofram influência. No entanto, o I de Moran global bivariado constatado foi ainda menor do que os anteriores, 0,1645, ainda que significativo, com um pseudo p -valor de 0,001, mas 135 municípios não alcançaram estatísticas LISA significantes. Os padrões encontrados foram semelhantes aos anteriores, sobretudo ressaltando relações diretas.

5.2. Modelo econométrico

Inicialmente, para a modelagem, foram escolhidas variáveis socioeconômicas consideradas relevantes para a explicação das taxas de suicídio, em conformidade com as recomendações da literatura empírica e teórica. À vista disso, foram determinados para a matriz de variáveis X : o índice de desenvolvimento, a renda e os níveis superiores de educação, que são vistos como elementos que levam à maior independência dos indivíduos e são utilizadas em praticamente todos os modelos da literatura empírica revisada, bem como também a desigualdade (GONÇALVES et al., 2011; DANTAS et al., 2018; SANTOS; BARBOSA, 2017; LOUREIRO et al., 2010; SHIKIDA et al., 2007; SCHNITMAN et al., 2011; FRAGA et al., 2016). A urbanização já explorada na AEDE, teorizada e verificada empiricamente por Durkheim (2000) e abordada também por Gonçalves et al. (2011), deverá constar. O desemprego, que espera-se ter uma relação inversa com a taxa de suicídio, também é verificado por Dantas et al. (2018), Santos e Barbosa (2017), Loureiro et al. (2010) e Fraga et al. (2016), e verificou-se a fecundidade, vista por Durkheim (2000) como um elemento que fortalece os laços familiares e que, portanto, tem relação inversa com a taxa de suicídio. No Quadro 1, pode ser observada a descrição das variáveis.

Com essas variáveis, foi estimado um modelo clássico por MQO com a intenção prévia de verificar os testes de detecção de relação espacial no modelo, haja vista que a literatura e a AEDE já indicam a presença desse fenômeno. Contudo, os testes sobre os resíduos rejeitaram as hipóteses de normalidade e homocedasticidade. Assim, foi adicionada uma variável *dummy* para caracterizar os municípios identificados como *outliers*. Os resultados dos testes desses modelos iniciais estão presentes na Tabela A1, no apêndice. Nota-se que a presença da *dummy* melhorou bastante o resultado dos testes de heterocedasticidade, mas não foi suficiente para não rejeitar a hipótese de normalidade.

Ainda assim, é dado continuidade à interpretação dos testes do modelo com a variável *dummy*, seguindo o procedimento proposto por Florax et al. (2003). O teste I de Moran para os resíduos foi significativo, demonstrando a existência de relações espaciais no modelo. Os testes LM para os termos de defasagem espacial do suicídio (ρ) e de erro espacial (λ) foram todos significantes, com exceção do teste robusto para λ , sendo o valor da estatística do teste LM_ρ sempre mais forte do que o LM_λ .

Por conseguinte, foi estimado o modelo SAR por dois métodos diferentes, o da Máxima Verossimilhança (MV) e o de Mínimos Quadrados em Dois Estágios (MQ2E) utilizando as defasagens espaciais das variáveis explicativas como Variáveis Instrumentais (VI), considerando que no MV exige-se o pressuposto de normalidade assintótica dos resíduos, enquanto que o método por VI essa hipótese não precisa ser satisfeita. O resultado de ambos encontra-se, respectivamente, nas Tabelas A2 e A3, no apêndice.

Ao nível de 5% de significância estatística, no modelo por MV, o índice de Gini e o IDHM foram significantes, ambos com relações diretas, isto é, quanto maior o desenvolvimento e a desigualdade, mais elevadas serão as taxas de suicídio. Por outro lado, a renda per capita apresentou uma relação significativa e inversa, na qual maiores níveis de renda podem levar a uma redução da taxa de suicídios, apesar da constatação de um baixo coeficiente em termos absolutos. No modelo por VI, considerado mais adequado por não supor normalidade, dos parâmetros do vetor $\hat{\beta}$, apenas o índice de Gini foi significativo e também positivo. Em ambos, o parâmetro de autocorrelação ρ foi

positivo e significativo. O teste Breuch-Pagan para o primeiro confirmou a homocedasticidade dos resíduos e o teste LR (Razão de Verossimilhança) ressaltou a significância do ρ . No modelo por VI, o teste Anselin-Kelejian comprovou que não há mais autocorrelação espacial nos resíduos do modelo.

Tabela 1 – Testes estatísticos do modelo estimado por MQO

Teste	Valor	Prob.
Jarque-Bera	4.8248	0.0896
Breusch-Pagan	14.6127	0.1021
Koenker-Bassett	17.6980	0.0388
<i>I</i> de Moran (ϵ)	2.4131 (I=0.1039)	0.0158
LM_ρ	10.6743	0.0011
LM_ρ * (robusto)	7.3737	0.0066
LM_λ	4.1379	0.0419
LM_λ * (robusto)	0.8373	0.3602
$LM(\rho + \lambda)$	11.5115	0.0032

Fonte: Elaboração própria.

Contudo, tendo em vista os problemas envolvidos nessa estimação inicial, sobretudo relacionados à normalidade dos resíduos, buscou-se alterar a especificação do modelo com a mudança de variáveis explicativas, visando ao alcance de melhores diagnósticos. Assim, a nova especificação trocou o índice de desigualdade de Gini pelo de Theil (usado por Schnitman et al. (2011)), as taxas de fecundidade e de desocupação foram excluídas e foram incluídas as variáveis razão de dependência (usado por Schnitman et al. (2011), Dantas *et al.* (2018) e Santos e Barbosa (2017)), pobreza (usado por Schnitman *et al.* (2011), Loureiro *et al.* (2010), Gonçalves et al. (2011), Fraga et al. (2016), Dantas et al. (2018) e Santos e Barbosa (2017)) e a o decil dos ricos, que ao determinar a renda mínima dos 10% mais ricos ressalta quão abastados são os ricos de uma sociedade, expondo também, certamente, o grau de independência desses indivíduos.

Os testes do novo modelo estimado por MQO estão presentes na Tabela 1. A nova especificação gerou testes que não rejeitam as hipóteses de normalidade e homocedasticidade dos resíduos (exceção do Koenker- Bassett). Um adendo deve ser feito também para a possibilidade de algum nível de multicolinearidade devido à semelhança de algumas variáveis, caráter bastante comum atribuído a esses tipos de dados socioeconômicos. Quanto aos testes de caráter espacial, o diagnóstico foi igual ao do modelo anterior, apenas com valores de estatísticas diferentes. Consequentemente, estimou-se um modelo SAR por Máxima Verossimilhança. Os resultados são observados na Tabela 2.

Os resultados da Tabela 2, possivelmente o modelo estimado mais adequado, constataram um ρ estatisticamente significativo ao nível de 5%, ressaltado ainda pelo teste *LR* que confirmou o modelo SAR como a especificação mais apropriada, comparada a um modelo sem defasagem espacial da variável dependente. Assim, pode ser afirmado, segundo o modelo da Tabela 2, com 5% de significância, que uma elevação no valor de uma unidade da taxa de suicídio na média das taxas das cidades vizinhas que dividem alguma fronteira (ou no aumento da taxa de cada uma dessas cidades em uma unidade) deve elevar, em média, a taxa de suicídio no município em questão.

Quanto às variáveis explicativas da matriz *X*, foi observado que o ensino superior tem uma relação direta, com significância aceita na margem, confirmando as formulações de Durkheim (2000). Além disso, a nova variável incorporada, razão de dependência, obteve significância e relação inversa com o suicídio, ressaltando que os suicídios ocorrem mais na população economicamente ativa antes de aposentar.

Tabela 2 – Resultados do modelo SAR

Variável	Coefficiente	Erro-padrão	Estatística Z	Prob.
Constante	4.1626	14.5161	0.2868	0.7743
Defasagem espacial (ρ)	0.2546	0.0841	3.0284	0.0025
Taxa de urbanização	-0.3095	1.6485	-0.1877	0.8511
IDH-M	21.8730	17.2905	1.2650	0.2059
Ensino superior	0.4733	0.2508	1.8873	0.0591
Renda <i>per capita</i>	0.0004	0.0137	0.0300	0.9761
Índice de Theil	8.0818	4.0591	1.9910	0.0465
Razão de dependência	-0.2215	0.0740	-2.9914	0.0028
Pobreza	-0.0049	0.0894	-0.0549	0.9562
Decil dos ricos	-0.0111	0.0048	-2.3205	0.0203
<i>Dummy outliers</i>	15.4698	1.2925	11.9685	0.0000

Teste de Breusch-Pagan: 15.4142 (Prob.= 0.08017); Pseudo- $R^2 = 0.5967$;

Teste de Razão de Verossimilhança (ρ): 9.2751 (Prob.= 0.0023).

Fonte: Elaboração própria.

A variável incorporada decil dos ricos apresentou significância estatística, mas com uma relação inesperada (como com a renda no primeiro modelo), já que se esperava que o maior nível de riqueza fosse levar a maiores taxas de suicídio, segundo Durkheim (2000). Já o índice de Theil tem uma relação direta e significativa com o suicídio, evidenciando que quanto maior a desigualdade de renda observada, maior serão os níveis de suicídio. Esses dois resultados devem estar em congruência com a confirmação da tese verificada e confirmada por Loureiro et al. (2010), de que a felicidade está atrelada ao consumo relativo e não ao consumo absoluto, ao observar que a desigualdade tem uma relação direta e a pobreza (e possivelmente a riqueza) têm uma relação inversa com o suicídio.

A urbanização não mostrou relevância em nenhum dos modelos estimados, provavelmente ressaltando que morar no campo ou em zonas urbanas não importa para as taxas de suicídio no Rio Grande do Norte. O IDHM, apesar de ter sido relevante no primeiro modelo, deixou de ser significativo na nova especificação. As influências da renda e da pobreza sobre o suicídio também não podem ser consideradas estatisticamente diferentes de zero. O grande destaque é para a defasagem espacial do suicídio, que se mostrou relevante em todos os modelos para explicar o suicídio, ressaltando o efeito contágio.

6. Conclusão

Esta pesquisa buscou verificar padrões espaciais da taxa de suicídio e suas relações com fatores socioeconômicos, partindo da tese enfatizada por Durkheim (2000), de que parte dos suicídios ocorrem devido a uma espécie de imitação, na qual indivíduos propensos a cometer suicídio teriam uma maior possibilidade de realizá-lo após o conhecimento de casos próximos. Dessa forma, o principal objetivo foi testar a hipótese para os municípios do Rio Grande do Norte para o que ficou caracterizado como um efeito contágio do suicídio.

Assim, a partir da análise espacial realizada, ficou constatado o efeito contágio entre os municípios do RN, apesar de indicadores de autocorrelação espacial serem considerados baixos e ressaltados em microrregiões específicas, mas foram sempre significantes, seja na Análise Exploratória de Dados Espaciais ou na modelagem econométrica-espacial.

A análise econométrica, embora tenha verificado relações espaciais sempre evidentes e significantes, não demonstrou essa constância e asserção com as variáveis socioeconômicas. A análise das significâncias estatísticas e a direção dos sinais se tornaram maleáveis e muito sensíveis às mudanças de especificação do modelo, certamente comprovando as conclusões de Rehkopf e

Buka (2006), que avaliam uma grande heterogeneidade de comportamentos dessas variáveis em estudos desse tipo. Além de ocorrer a inexistência de um modelo teórico fechado sobre o assunto, provavelmente por sua impossibilidade de formulação.

Dessa forma, uma análise econométrica mais profunda, considerando outras especificações, áreas geográficas, métodos e tipos de variáveis, embasando-se em uma vasta revisão da literatura, seria recomendada, ficando como sugestão para futuras pesquisas. Como exemplo, as relações possíveis envolvendo o interior do estado, local no qual ficou constatado que municípios mais distantes do litoral apresentam maiores taxas de suicídio (relações com a incidência solar, produção agropecuária, pluviometria, gastos com saúde etc.).

Portanto, as inferências feitas a partir das variáveis socioeconômicas têm como resultado poucas conclusões, mas ainda permitem que sejam propostas políticas de prevenção ao suicídio. A teoria indica que a população com alto grau de independência socioeconômica é mais propensa ao suicídio e os resultados mostram que isso ocorre sobretudo com aqueles com maior nível de escolaridade, nomeadamente, com ensino superior completo. Portanto, parcela das políticas de prevenção ao suicídio poderiam ser aplicadas sobre essa parte da população. Também há uma indicação, contrariamente ao que se poderia pensar, que suicídios podem ser mais frequentes sobre a população economicamente ativa relativamente aos dependentes mais idosos.

Além disso, a relação direta entre taxa de suicídio e desigualdade de renda mostra que a redução do suicídio pode ser um subproduto de políticas de redistribuição de renda. Com relação aos resultados puramente espaciais, sua interpretação pode levar a sugestão de políticas localmente focadas, direcionadas para áreas interioranas e rurais, como nas microrregiões do Seridó, de Pau dos Ferros e de Umarizal que apresentam padrões de altas taxas de suicídio com evidentes efeitos de contágio.


Referências

- ANSELIN, L. Local indicators of spatial association - LISA. *Geographical Analysis*, v. 27, n. 2, p. 93-115, 1995.
- ANSELIN, L. Exploring spatial data with GeoDaTM: a workbook. *Center for Spatially Integrated Social Science*. Urbana, IL. 2005.
- ANSELIN, L.; REY, S. J. *Modern spatial econometrics in practice: a guide to GeoDa, GeoDaSpace and PySAL*. GeoDa Press LLC, 2014.
- ANSELIN, L.; SYABRI, I.; SMIRNOV, O. *Visualizing multivariate spatial correlation with dynamically linked windows*. In: Proceedings of the CSISS Workshop on New Tools for Spatial Data Analysis. Santa Barbara, CA. 2002.
- ASSUNÇÃO, R. M. Estatística espacial com aplicações em epidemiologia, economia e sociologia. In: Congresso da Associação Brasileira de Estatística. *Anais...* São Carlos: Associação Brasileira de Estatística, 2001.
- CRUZ, B. D. O.; FURTADO, B. A.; MONASTERIO, L.; JÚNIOR, R. *Economia regional e urbana: teorias e métodos com ênfase no Brasil*. Brasília: IPEA, 2011.
- CUTLER, David M.; GLAESER, E. L.; NORBERG, K. E. Explaining the rise in youth suicide. In: GRUBER, J. (Ed.) *Risky behavior among youths: an economic analysis*. University of Chicago Press, 2001. p. 219-270.
- DANTAS, A. P.; AZEVEDO, U. N. D.; NUNES, A. D.; AMADOR, A. E.; MARQUES, M. V.; BARBOSA, I. R. Analysis of suicide mortality in Brazil: spatial distribution and socioeconomic context. *Revista Brasileira de Psiquiatria*, v. 40, n. 1, p. 12-18, 2018.

- DURKHEIM, E. O. *O suicídio: estudo de sociologia*. Tradução de Mônica Stahel. São Paulo: Martins Fontes, 2000.
- DUTRA, E.; ROEHE, M. V. Suicídio de agricultores no Rio Grande do Norte: compreensão fenomenológica preliminar dos aspectos existenciais. *Psicologia em Revista*, v. 19, n. 1, p. 108-118, 2013.
- FLORAX, R. J. G. M.; FOLMER, H.; REY, S. J. Specification searches in spatial econometrics: the relevance of Hendry's methodology. *Regional Science and Urban Economics*, v. 33, n. 5, p. 557-579, 2003.
- FRAGA, W. S.; MASSUQUETTI, A.; GODOY, M. R. Determinantes socioeconômicos do suicídio: um estudo para os municípios do Brasil e do Rio Grande do Sul. *Revista Econômica*, v. 18, n. 2, p. 1-37, 2016.
- GONÇALVES, L. R. C.; GONÇALVES, E.; OLIVEIRA JÚNIOR, L. B. Determinantes espaciais e socioeconômicos do suicídio no Brasil: uma abordagem regional. *Nova Economia*, v. 21, n. 2, p. 281-316, 2011.
- LESAGE, J. P.; PACE, R. K. *Introduction to spatial econometrics*. Statistics: textbooks and monographs. New York: CRC Press, 2009.
- LOUREIRO, P. R. A.; MENDONÇA, M. J. C.; SACHSIDA, A. *Os determinantes econômicos do suicídio: um estudo para o Brasil*. Rio de Janeiro: Instituto de Pesquisa Econômica Aplicada (IPEA), 2010. (Texto para Discussão, n. 1487).
- MONTEIRO, A. M. V., CÂMARA, G.; CARVALHO, M. S.; DRUCK, S. *Análise espacial de dados geográficos*. Planaltina: Embrapa, 2004.
- REHKOPF, D. H.; BUKA, S. L. The association between suicide and the socio-economic characteristics of geographical areas: a systematic review. *Psychological Medicine*, v. 36, n. 2, p. 145-157, 2006.
- SANTOS, E. G. O; BARBOSA, I. R. Conglomerados espaciais da mortalidade por suicídio no nordeste do Brasil e sua relação com indicadores socioeconômicos. *Cadernos Saúde Coletiva*, v. 25, n. 3, p. 371-378, 2017.
- SHIKIDA, C. D.; ARAUJO JR, A. F.; GAZZI, R. A. V. Teoria econômica do suicídio: estudo empírico para o Brasil. *Análise Econômica*, v. 25, n. 48, p. 123-147, 2007.
- SCHNITMAN, G.; KITAOKA, E. G.; AROUCA, G. S. D. S.; LIRA, A. L. D. S.; NOGUEIRA, D.; DUARTE, M. B. Taxa de mortalidade por suicídio e indicadores socioeconômicos nas capitais brasileiras. *Revista Baiana de Saúde Pública*, v. 34, n. 1, p. 44-59, 2011.
- SGOBIN, S. M. T. *Custos diretos e indiretos de tentativas de suicídio de alta letalidade internadas em hospital geral*. Dissertação (Mestrado em Saúde Pública). Faculdade de Ciências Médicas da Universidade Estadual de Campinas (UNICAMP). Campinas, 2013.

ORCID

Stefany Silva Amaral  <https://orcid.org/0000-0002-6420-4964>

 Este artigo está licenciado com uma Licença Creative Commons Atribuição 4.0 Internacional.

Apêndice

Tabela A1 – Testes estatísticos do modelo inicial estimado por MQO, com e sem *dummy*

Teste	Modelo inicial		Modelo c/ <i>Dummy</i>	
	Valor	Prob.	Valor	Prob.
Jarque-Bera	89.7920	0.0000	8.6616	0.0132
Breusch-Pagan	39.6330	0.0000	9.8858	0.2731
Koenker-Bassett	17.2520	0.0158	11.1090	0.1956
<i>I</i> de Moran ($\hat{\epsilon}$)	1.9820 (I=0.082)	0.0475	3.1814 (I=0.1424)	0.0015
LM_ρ	7.7340	0.0054	17.3620	0.0000
LM_ρ * (robusto)	12.9810	0.0003	11.0117	0.0009
LM_λ	2.6310	0.1048	7.7669	0.0053
LM_λ * (robusto)	7.8780	0.0050	1.4165	0.2340
$LM(\rho + \lambda)$	15.6120	0.0004	18.7786	0.0001

Fonte: Elaboração própria.

Tabela A2 – Resultados do modelo SAR inicial estimado por Máxima Verossimilhança

Variável	Coefficiente	Erro-padrão	Estatística Z	Prob.
Constante	-17.8946	10.3855	-1.7230	0.0849
Defasagem espacial (ρ)	0.3254	0.0838	3.8826	0.0001
Taxa de urbanização	0.8466	1.7310	0.4891	0.6248
IDH-M	33.6046	16.3913	2.0502	0.0404
Renda <i>per capita</i>	-0.0147	0.0070	-2.0915	0.0365
Taxa de desocupação	-0.0236	0.0703	-0.3356	0.7372
Taxa de fecundidade	-1.1036	0.8272	-1.3341	0.1822
Índice de Gini	14.5413	6.5692	2.2136	0.0269
Ensino superior	0.1587	0.2402	0.6608	0.5087
<i>Dummy outliers</i>	15.5278	1.3435	11.5576	0.0000

Teste de Breusch-Pagan: 11.4714 (Prob.= 0.1763); Pseudo- $R^2 = 0.5663$;

Teste de Razão de Verossimilhança (ρ): 14.7594 (Prob.= 0.0001).

Fonte: Elaboração própria.

Tabela A3 – Resultados do modelo SAR inicial estimado por MQ2E (defasagem espacial das variáveis explicativas (*WX*) como variáveis instrumentais de *Wy*)

Variável	Coefficiente	Erro-padrão	Estatística Z	Prob.
Constante	-17.2005	10.3142	-1.6677	0.0954
Defasagem espacial (ρ)	0.5100	0.1243	4.1037	0.0000
Taxa de urbanização	0.4040	1.5605	0.2589	0.7957
IDH-M	26.0741	15.3502	1.6986	0.0894
Renda <i>per capita</i>	-0.0105	0.0073	-1.4467	0.1480
Taxa de desocupação	0.0009	0.0796	0.0118	0.9906
Taxa de fecundidade	-0.5617	0.8179	-0.6867	0.4922
Índice de Gini	15.3549	7.2821	2.1086	0.0350
Ensino superior	0.1149	0.2694	0.4264	0.6698
<i>Dummy outliers</i>	15.5125	1.5668	9.9009	0.0000

Teste de Anselin-Kelejian: 3.3970 (Prob.= 0.0653); Pseudo- $R^2 = 0.5683$.

Fonte: Elaboração própria.